Center Reflections

A monthly publication highlighting activities at the W.M. Keck Foundation Center for Molecular Structure

California State University Fullerton

Volume 1, Issue 7 November/December 1999

X-RAY ABSORPTION AND DIFFRACTION STUDIES OF THE METAL TO INSULATOR TRANSITION IN $NiS_{1.52}Se_{0.48}$

San Jose State University

In Professor Juana Acrivos' laboratory at SJSU, studies of $NiS_{1.52}Se_{0.48}$ have focused on understanding the nature of the compound's temperature dependent metal to insulator transition, which exists at Se compositions varying between x~0.4 and x~0.6 and in the 0K to 115K temperature range. The X-ray Absorption Spectrum (XAS) at the Ni and K alpha edges in the 4K to 150K temperature range was measured. The changes in the XAS over the temperature range correspond to the metal to insulator transition and support the Mott-Huddard model for metal to insulator transitions proposed for these compounds.

X-ray diffraction studies on this compound were conducted at CMolS by Thaddeus Norman, a graduate student in the Acrivos laboratory. The compound crystallizes in a cubic system (Figure 1) with unit cell dimensions a=5.7539 Å. Earlier work established that $NiS_{2-x}Se_x$ compounds obey Vegard's Law. According to this principle the unit cell dimensions of a mixed crystal such as this vary linearly with composition.

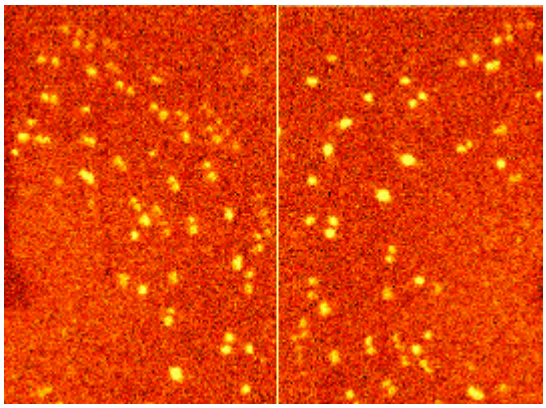


Figure 1: X-Ray Diffraction Pattern for $NiS_{2-x}Se_x$.

From a Vegard's Law plot (Figure 2) the value of x was determined to be 0.48. The unit cell dimensions were used to construct the crystal system for $NiS_{1.52}Se_{0.48}$, and this model was used for the initial guess for the EXAFS fitting routine. Since both NiS_2 and $NiSe_2$ crystals have Pa3 symmetry the Pa3 space group was used to generate the x, y, z positions for the atoms in the lattice. X-ray diffraction data on NiS_2 and $NiSe_2$ reveal that Ni occupies the 4 a position in the lattice and the chalcogenide atoms occupy the 8 c positions. The distance between the atoms was then calculated using the equation

$$d = [(xa-xb)^2 + (ya-yb)^2 + (za-zb)^2]^{1/2}$$

where d is the distance between the atoms.

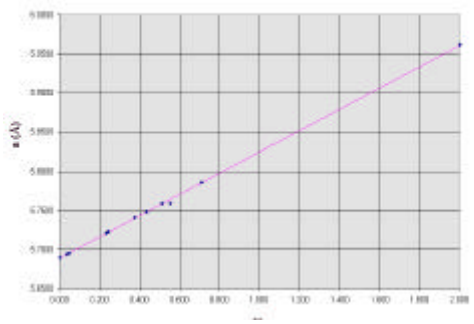


Figure 2: Vegard's Law plot for $NiS_{2-x}Se_x$. The equation for the line is y=mx+b where m=0.1359, b=5.6890 and R^2=0.999534. In the early 20th century Vegard demonstrated that for many compound unit cell dimensions are linearly dependent on composition.

In the unit cell of $NiS_{1.52}Se_{0.48}$ there are four Ni atoms. In the model compound used for the Ni edge each Ni atom is surrounded by three Ni atoms at 4.0634Å distance, and eight chalcogenide atoms two at 2.4196Å, two at 3.9098Å, two at 5.6390Å and two at 7.5988Å. The unit cell from the perspective of the chalcogenide atoms is more complicated. Each chalcogenide atoms does not have the same number of atoms at the same distance. So for simplicity the model position for the Se edge was based on the atom which had three Ni atoms at the closest distance. For the initial guess only one Se atoms was used at the 2.4196Å distance and none at the other distances.

Thaddeus Norman holds both BS and Masters degrees from San Jose State University. His career objective is a Ph.D. Thaddeus is currently employed in the silicon valley in the area of solid state chemistry.

Professor Juana Acrivos received her *D. Sc.* from the University of Habana and her Ph.D. with minor in Mathematics, in Physical Chemistry, from the University of Minnesota. She is the recipient of many international fellowships and awards and is a member of the New York Academy of Sciences. For more about her research, check out http://www.sjsu.edu/testupdates/faculty/Acrivos/.

Structural Genomics: A Key Underpinning for Biomedical Research in the 21st Century

In April 1999, the Merck Genome Research Institute and the National Cancer Institute's Cancer Genome Anatomy Project (both support structural genomics technology development and research) brought together 27 scientists working at the forefront of a new field. This new vision of comprehensive molecular analysis, in which functions of individual genes will be studied within the context of all the genes in an organism or a cell, is already being established as a key underpinning for research in the 21st century. Captured under the broad term "structural genomics", this vision includes related subcategories of transcriptome, proteome and physiome.

Structural genomics has been driven in large part by recent scientific advances. Large-scale application of expressed sequence tag technology has resulted in the identification of tens of thousands of human genes. Over the past few years the genomic sequences of various microbes, including model systems such as *E. coli* and *S. cerevisiae*, as well as pathogens such as *H. pylori* and *M. tuberculosis* have been determined. Also, the essentially complete sequence of the multicellular eukaryote, *C. elegans*, has recently been learned. Computing technology and bioinformatics methodologies have allowed us to manage enormous databases and to rationally mine them. Analysis of genomic sequences has revealed

previously unknown genes. We can now study a particular gene with knowledge of potential partners in the genetic network and put together a complete picture of how the genome performs its functions. By comparing protein sequences, and through prediction of secondary and tertiary structure, we can interpret existing functional data, design ligands and construct mutants for testing new functional hypotheses. The task before us is to build complete catalogs of genomes and their products and to learn how these macromolecules interface dynamically to produce complex cells and organisms.

Two great challenges in structural genomics are to assign potential protein function and to understand what proteins may perform related activities. Homologous proteins may be identified based on translations in DNA sequences and comparisons of primary amino acid sequences. However, similar protein structures can be derived from primary sequences that are quite dissimilar (myoglobin and hemoglobin are the first structural examples of this).

Many laboratories are engaged in the development of structural genomics technologies that will be broadly available as research tools to the entire research community. The focus is to develop assays and methodologies with improved accuracy and speed. Efforts include:

- cDNA cloning/sequencing technology
- bioinformatics

e.g. - algorithms to predict gene/protein function based on sequence or to predict protein folds

- *ultra- high-throughput* proteomics

identifying proteins, their functional states, post-translational modifications, quantity, rate of synthesis and turnover, gene expression and protein purification, crystallization and x-ray crystallographic analysis.

As of April 1999, there were more than 3.5 million sequences in GenBank and more than 2.6 million bases from more than 41,000 species, including 301 megabases of finished human sequence. There are more than 10,000 atomic coordinate files in the Protein Data Bank. Thus, we have an enormous task at hand to manage, utilize and learn from all the information being generated. Perhaps the most challenging aspect of science will be related to bioinformatics and how we will capture the information in these databases. We will need to find new ways to interact, collaborate, generate and share information. Without sharing fundamental tools, knowledge and expertise, we will have great difficulty harnessing this structural revolution. Macromolecular crystallography will have a significant role to play.

Homology Modeling of Aminoglycoside 6'-N-Acetyltransferase from Klebsiella pneumoniae AAC(6')-Ib

CSU Fullerton

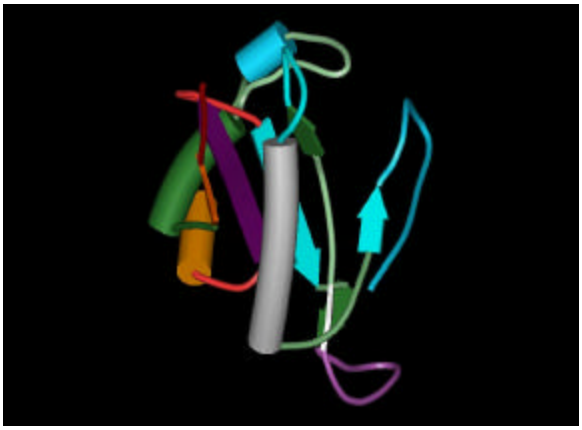
According to the American Medical Association, the global increase in resistance to antimicrobial drugs, including the emergence of bacterial strains that are resistant to all available antibacterial agents, has created a public health problem of potentially crisis proportions. The success of antibiotics accounts in part for the resistance problem. The U.S. Centers for Disease Control and Prevention (CDC) estimates that one-third of the 150 million antibiotic prescriptions written for outpatients every year are unnecessary. Such misuse of drugs is a major factor in the 50 percent jump in antibiotic-resistant infections over the past 23 years, the CDC says-- including the rise of bacteria resistant to all known drugs. The pharma-

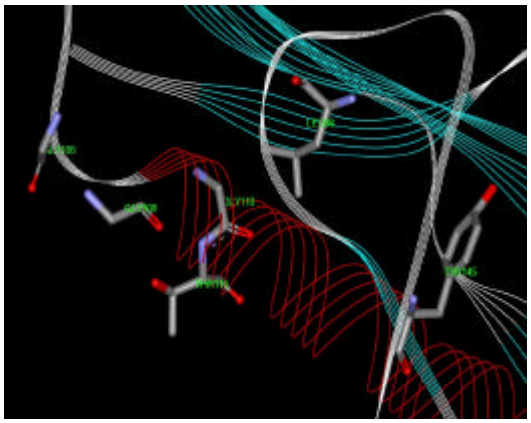
ceutical industry is losing its race against developing resistance by pathogens, and the costs to combat drug resistance are increasing.

One approach to combating drug resistance is to study the molecular mechanisms of bacterial resistance to antibiotics and use the knowledge gained to design rational strategies to overcome the problem. Using this approach, antibiotics refractory to modifying enzymes, such as amikacin, vancomycin, Synercid and Zyvox have been developed. However, bacterial resistances to amikacin and vancomycin have rapidly evolved.

Aminoglycoside N-acetyltransferases are aminoglycoside-resistant enzymes in bacteria. Some of these enzymes can modify the clinically important aminoglycosides that have been used as antibiotics. Professors Marcelo Tolmasky (Biology) and Katherine Kantardjieff (Chemistry) are determining the crystal structure of the aminoglycoside 6'-N-acetyltransferase from clinical strains of *Klebsiella pneumoniae* (AAC(6')-Ib). This enzyme imparts resistance to amikacin and other aminoglycoside drugs by catalyzing the acetylation of the molecules using acetyl-CoA. At this time, overexpression of the enzyme in *E. coli* has produced a) improperly folded protein with no activity; b) protein with limited solubility.

Last year, the first and the only available crystal structure of an enzyme in the N-acetyltransferase family (NAT), aminoglycoside 3'-N-acetyl transferase from *Serratia marcescens*, SmAAC(3'), was reported by Stephen K. Burley. This structure has provided useful data to predict a model of the AAC(6')-Ib enzyme. Goragot Wisedchaisri, a graduate student from Thailand in Kantardjieff's laboratory, has conducted sequence analysis of the primary structure of AAC(6')-Ib, predicted the secondary structure using several available programs and databases, and completed homology modeling using the related SmAAC(3') as a template. Goragot's schematic model of the AAC(6')-Ib enzyme is shown above. The model contains four α-helices and six β-sheet strands with nine loops. The predicted model for the acetyl-CoA binding site of AAC(6')-Ib is shown below. Amino acids shown as stick model are identical to amino acids interacting with CoA in crystal structure of SmAAC(3').





Websites of Interest

- ***Journal of the Chemical Computing Group, Inc.*** http://www.chemcomp.com/ Chemical Computing Group Inc. develops and markets high-end scientific software and services for High Throughput Screening and Computer Aided Molecular Design applied to Life and Materials Sciences.
- ***CRYSTOOL*** http://www-structure.llnl.gov/crystool/crystool.htm See Volume 1, Issue 6 of *Center Reflections*.
- ***Growing Crystals that Will Make Your Crystallographer Happy*** http://rocket.chem.ualberta.ca/xray/links.html. See Volume 1, Issue 6 of *Center Reflections*
- ***Crystallography 101*** (www-structure.llnl.gov/Xray/101index.html) - See Volume 1, Issue 3 of *Center Reflections*.

Upcoming Events

November 20, 1999: **Southern California Council on Undergraduate Research Annual Conference,** Loyola Marymount University, Los Angeles, CA.

January 13-14, 2000: **12th Annual CSU Biotechnology Symposium,** CalPoly Pomona Kellogg Conference Center.

February 12-16, 2000: **Biophysical Society** Annual Meeting, New Orleans, LA. http://www.biophysics.org/biophys/society/annmtg/

March 26-30, 2000: **American Chemical Society** National Meeting, San Francisco, CA. http://www.acs.org/meetings/sanfran2000/

April 15-18, 2000: **Experimental Biology 2000 FASEB** Meeting and Scientific Exposition, San Diego, CA. http://www.faseb.org/eb2000

July 22-27, 2000: **American Crystallographic Society** Annual Meeting, St. Paul, MN. http://nexus.hwi.buffalo.edu/ACA/ACA-Annual/StPaul/StPaul.html

W.M. Keck Foundation Center for Molecular Structure

Department of Chemistry and Biochemistry

California State University Fullerton

800 N. State College Blvd.

Fullerton, CA 92831

http://www-structure.llnl.gov/scaurcon99/cmols2.html

Director: Dr. Katherine Kantardjieff

kkantardjieff@fullerton.edu

Staff Scientist: We're looking for someone. Contact us!